

United States Patent Application

Title of the Invention

NETWORK RELAYING APPARATUS AND NETWORK
RELAYING METHOD CAPABLE OF HIGH-SPEED
ROUTING AND PACKET TRANSFER

Inventors

Kazuo SUGAI,
Takeshi AIMOTO,
Takeki YAZAKI,
Nobuhito MATSUYAMA,
Yoshihito SAKO,
Tomohiko TANABE.

- 1 -

NETWORK RELAYING APPARATUS AND NETWORK RELAYING
METHOD CAPABLE OF HIGH-SPEED ROUTING AND PACKET TRANSFER

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of three applications U.S. Serial Nos. _____ being filed by the same applicants as those of this application, based on Japanese patent application Nos. 11-045959, 11-046837 and 11-046579 filed on February 24, 1999, respectively and assigned to the present assignee. The contents of these application are incorporated by reference herein.

BACKGROUND OF THE INVENTION

The present invention relates to a network relaying apparatus and a network relaying method, or in particular to a network relaying apparatus including a
5 router of a computer network system which is capable of searching at high speed for a destination of a packet input and a network relaying search method.

Generally, in a network system, a network relaying apparatus such as a router or a bridge is used
10 for connecting a plurality of networks. The router checks the destination address of a packet received from a network or a subnet connected, determines the destination of the packet, and transfers the packet to a network or a subnet which is connected with the
15 destination router or host.

Fig. 13 is a diagram showing a configuration of a conventional network relaying apparatus. In Fig. 13, a router 100 includes a routing manager (RM) 110, router buses 120, network interfaces (NIF) 130 and ports 140. Each port 140 is connected to an appropriate network 150.

Each network interface 130 receives a packet from a network connected to the port 140, and transmits the received packet through the router bus 120 to the routing manager 110. The routing manager 110 includes a routing table for holding the routing information, and using this routing information, determines the network 150 of the destination from the address of the packet received, and transmits the packet to the network interface 130 of the port 140 connected to the network 150. The network interface 130 that has received the packet from the routing manager 110 sends out the packet to the destination network 150. The routing manager 110 updates and maintains the routing information held in the routing table based on the header information of the packet received, and has the function of overall management of the router 100.

An explanation will be given of the route search process for searching for a port outputting the next address to which the packet is to be transferred upon receipt of the packet and outputting the packet. Normally, the route search uses a route search table (routing table) prepared from the component definition

information and the information obtained by exchange
between the routers. The routing table is for searching
the information (next hop information) as to the output
port, the next hop address and whether the network is
5 directly connected or not with a set of the network
address and the network mask length as a key.

As another conventional system, JP-A-05-199230
(US Patent Serial No. 5,434,863) discloses an
internetwork system and a communication network system
10 which can flexibly meet the size requirement of the
network without adversely affecting the high-speed
routing process. In these systems, a router manager and
a plurality of routing accelerator modules are coupled
to each other with a high-speed bus. Also, each routing
15 accelerator is connected with a plurality of independent
communication ports. In these conventional systems, a
plurality of the routing accelerators makes possible a
high-speed routing and by adding the routing accel-
erators, the requirement for increasing the network size
20 can be easily met.

SUMMARY OF THE INVENTION

In recent years, the demand has increased for
the dynamic routing in which the relaying information
for routing is dynamically generated, added, changed or
25 deleted by recognizing the configuration of the network
in operation. Specifically, the router requires the
processing of the routing protocol (such as the Routing

Information Protocol (RIP) or Open Shortest Path First (OSPF) included in TCP/IP protocols) for exchanging information on the network between the routers. Further, the processing of the network management
5 protocol (such as Simple Network Management Protocol (SNMP) which is one of the TCP/IP protocols) for communication of the management information such as the performance of the router with a management master station on the network is performed unavoidably by the
10 routing means in the prior art. This makes it impossible for the router to exhibit the relaying performance sufficiently. The conventional router, therefore, cannot easily meet the requirement of the high-speed lines such as the high-speed LAN (Local Area Network),
15 the wide-band ISDN and ATM (Asynchronous Transfer Mode) that have recently found practical applications.

Also, with the recent increase in the operating speed of the network, the data processing system used for the routers and bridges require a high
20 processing speed of the network controller searching for a transfer destination route from a memory. Further, considering the processor contention for memory access, the conventional routers limit the number of network controllers to absorb the reduced performance caused by
25 the memory access contention or, though low in cost effectiveness, unavoidably use a high-speed memory or a dual-port memory accessible from the processors or the network controller asynchronously. Also, the route

search in the conventional router is carried out mainly in software, thereby making high speed execution of the routing process difficult.

An object of the present invention is to
5 provide a network relaying apparatus and method for high speed routing while assuring a high communication quality (QoS), a high reliability and security.

Another object of the invention is to provide a network relaying apparatus and method for high speed
10 packet routing and packet transfer by executing the hardware processing for each function block including a transfer engine and a search engine.

Still another object of the invention is to provide a network relaying apparatus and method for
15 high-speed routing by dividing the routing process into the receiving process, the transmission process, the input search process and the output search process with required tables used independently of each other.

Yet another object of the invention is to
20 provide a relaying apparatus and method for realizing a higher speed by executing each process by pipelining.

Other objects, features and advantages of the present invention will become apparent from the following description of the embodiments of the
25 invention taken in conjunction with the accompanying drawings.

According to one aspect of the invention, there is provided a network relaying apparatus

connecting a plurality of networks for outputting the packets input from the networks, to the next transfer destination based on the route information, comprising:

at least a network interface connected to the
5 networks for controlling the interface with the networks;

at least a routing processor connected to one or a plurality of the network interfaces for routing the packets input from the network interfaces;

10 a routing manager for managing the internal components of the system; and

a connector for connecting the routing manager and each of a plurality of the routing processors;

wherein the routing processors each include:

15 a packet buffer for storing an input packet;

a high-speed readable and writable header memory accessible asynchronously with the packet buffer and adapted for storing the header information including the header and the internal header of the input packet;

20 a route table for storing the route information including the IP address of the next router corresponding to the destination internet protocol (IP) address;

an address search table for storing a media
25 access control (MAC) address of the next router corresponding to the IP address of the next router;

a flow search table for storing the action corresponding to the reference conditions including the

IP headers of the source and the destination;

a transfer engine for performing a receiving process for storing an input packet received from a network or the connector in the packet buffer, adding
5 the internal header to the packet header and storing the resulting header information in the header memory, and a transmission process for reading the input packet from the packet buffer, producing an output packet from the input packet stored in the packet buffer and the header
10 information stored in the header memory, and outputting the output packet to the connector or the network; and

a search engine for performing an input search process for searching the transfer destination information with reference to the route table based on the
15 header information stored in the header memory, and an output search process for searching the MAC address of the next router with reference to the address search table based on the IP address of the next router determined in the input search process and searching
20 various action including QoS with reference to a flow search table.

According to another aspect of the invention, there is provided a network relaying method for outputting the packets input from the networks, to a
25 transfer destination in a network relaying apparatus comprising at least a network interface connected to the networks, at least a routing processor for routing the packet input from the network interface, a routing

manager for managing the internal parts of the system,
and a connector for connecting the routing manager and
each of a plurality of the routing processors,
comprising:

- 5 a receiving process for storing the input
packet, and storing the header information separately
from the input packet by adding the internal header
including the input and output port numbers and the QoS
control information to the MAC header and the IP
10 (internet protocol) header of the input packet;
 an input search process for extracting the
destination IP address in the IP header from the header
information stored by the receiving process, and
searching the transfer destination information including
15 the IP address of the next router based on the
destination IP address;
 an output search process for extracting the IP
address of the next router determined by the input
search process, searching the MAC (media access control)
20 address of the next router based on the IP address,
searching the action information including the QoS based
on the reference conditions including the transfer
destination information and the destination information,
and storing the searched transfer destination
25 information and the action information in the header
information; and
 a transmission process for producing an output
packet based on the input packet and the header

information and outputting the output packet to the connector or the network interface.

Other objects, features and advantages of the present invention will become apparent from the following description of the embodiments of the invention taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a diagram showing a configuration of a network relaying apparatus according to the present invention.

Fig. 2 is a diagram showing an internal structure of a routing processor for explaining the operation of a network relaying apparatus.

Fig. 3 is a sequence diagram showing an outline of the operation of the network relaying apparatus.

Figs. 4A-4B are diagrams for explaining a packet buffer and a header RAM.

Figs. 5A-5C are diagrams for explaining each table used for route search.

Fig. 6 is a diagram for explaining the high-speed processing in the routing processor.

Fig. 7 is a diagram showing a configuration of a search engine in hardware.

Fig. 8 is a diagram for explaining the high-speed processing by pipelining control.

Fig. 9 is a diagram for explaining the flow search processing.

Fig. 10 is a diagram for explaining a flow search table.

5 Fig. 11 is a diagram for explaining a first input line limiting system.

Fig. 12 is a diagram for explaining a second input line limiting system.

10 Fig. 13 is a diagram showing a configuration of a conventional network relaying apparatus.

DESCRIPTION OF THE EMBODIMENTS

Detailed description of embodiments of the invention is made with reference to the drawings.

15 Fig. 1 is a diagram showing a configuration of a network relaying apparatus according to this invention. A router 1 includes a plurality of routing processors (RP) 10, a crossbar switch (CS) 20, at least a network interface (NIF) 30, at least a port 40, a routing manager (RM) 60 and a power supply (PS) 70.
20 Each port 40 is connected to an appropriate network. The network 50 is a LAN, a WAN or an ATM, for example. For assuring an improved reliability of the apparatus, the power supply 70 or each common part can be doubled as required.

25 The routing manager function is divided into the routing processors 10 for executing the routing function and the routing manager 60 for managing the

router 1. Further, the router 1 includes a plurality of routing processors 10 each having one or a plurality of network interfaces 30. The routing manager 60 has the function of overall management of the router 1 and at
5 the same time executes the route calculation function. Further, the routing manager 60 exchanges the routing information with other routers and distributes the routing information to each routing processor 10 within each router. The routing manager 60 has a dual
10 structure. The switch 20 has a crossbar switch or the like for communication and exchange between the routing processors 10 or between a routing processor 10 and the routing manager 60. The switch 20 is also formed in dual structure in the case under consideration. The
15 switch 20 may be replaced with a bus or the like for connection. Also, in the case where the crossbar switch is used, the connection route is not occupied by the routing manager 60 and one of the routing processors 10 but can be shared by a plurality of the routing
20 processors 10 at the same time.

Each routing processor 10 transfers packets through the network interface 30 connected thereto. A given routing processor 10 can also transfer a packet to the network 50 connected to another routing processor 10
25 through the switch 20. The routing processors 10 have each function thereof designed to perform a high-speed operation. More specifically, the routing processors 10 have such functions as switching, route search, forward-

ing, filtering, offering QoS and IP multicasting. Each routing processor 10 has an appropriate input buffer and an output buffer for each port 40 of the network interface 30 within it or for each of the other routing processors 10 and the routing manager 60. Each network interface 30 has one or a plurality of ports 40 for controlling the interface between the networks 50 and the routing processors 10.

Fig. 2 is a diagram showing the internal structure of the routing processor for explaining the operation of the network relaying apparatus. With reference to this diagram showing the internal structure of the routing processor 10, an explanation will be given of the operation of searching the route and transferring packets to the destination determined as the result of the route search.

The routing processors 10 each includes a transfer engine 13, a search engine 14, a header RAM 11, a packet buffer 12, a route table 15, an ARP (address resolution protocol) table 16, and a filter/QoS (flow search table) 17. The transfer engine 13 performs the packet input/output processing, for example. The search engine 14 mainly performs the route search and the flow search such as the QoS control based on the header information of the packet. The search engine 14 is configured with an exclusive LSI or the like hardware capable of high-speed processing.

The packet buffer 12 has the packet stored

therein until the transfer engine 13 transfers the input packet to the routing processor 10. The header RAM 11 extracts and stores only the header of the input packet. The header RAM 11 is configured of a memory having a
5 high read/write speed. This embodiment, in addition to the buffer memory for storing the packets received from the network or transferred from other data processing systems, comprises a header RAM 11 accessible asynchronously with the packet buffer 12. Thus, the
10 packet is stored in the packet buffer 12 while storing (copying) the header of the packet in the RAM 11 at the same time. Each processor of the transfer engine 13 and the search engine 14 fetches the header of the packet by use of the header RAM 11, and while analyzing the
15 header, the operation of reading/writing of the packet from or into the packet buffer 12 becomes possible. In this way, the header analysis of a packet and the transfer of other packets can be concurrently performed.

As long as the header information is being
20 read from the header RAM 11 by the search engine 14, the packet buffer 12 is not used by the processor. Therefore, the transfer engine becomes accessible to the packet buffer 12 for transmission or transfer thereby avoiding the competition for access to the packet buffer
25 12 between the search engine 14 and the transfer engine 13. The area for storing the packets received from the network and the header thereof can be configured separately from the area for storing the packets

transferred from the switch 20 and the header thereof.
This isolated configuration facilitates the packet
control.

The route table 15, the ARP table 16 and the
5 filter/QoS table are configured independently of each
other. Thus, the search engine 14 can access the tables
individually for read or write operation, thereby making
it possible to search for the routing information and
the QoS at high speed. Further, in order to realize the
10 high-speed routing, the pipelining process can be
carried out. Each table and the pipelining process will
be described in detail later.

Fig. 3 is a sequence diagram showing an
outline of the operation of a network relaying
15 apparatus.

First, when a packet is input to a first
network interface 30 through the network through a port,
the first network interface 30 transmits it to the
transfer engine 13. The transfer engine 13 stores the
20 received packet in the packet buffer 12. Also, the
transfer engine 13 extracts only the header of the input
packet and by adding the internal header, forms header
information, which is stored in the header RAM 11. The
internal header will be described later.

25 The search engine 14 reads the header infor-
mation by accessing the header RAM 11. Alternatively,
the transfer engine 13 may transfer the header
information stored in the header RAM 11 to the search

engine 14. In the search engine 14, the number or address of the router, the RP and the port of the destination, the information on the next transfer route such as a MAC (media access control) address and the
5 information for controlling the communication quality such as the QoS control information are searched for appropriately in accordance with the header information. The search engine 14 writes the destination information including the number or address searched and the
10 transfer control information including the action information such as the QoS information in the header RAM 11. The search engine 14 may alternatively transmit the transfer control information to the transfer engine 13.

15 In the transfer engine 13, an output packet is produced based on the packet stored in the packet buffer 12 and the header information (including the transfer control information) stored in the header RAM 11. The transfer engine 13 outputs the output packet thus
20 produced to the destination. In the the case where the transfer route is associated with any other routing processor 10, the transfer engine 13 sets the packet in queue for the buffer of the particular other routing processor 10, while in the case where the transfer route
25 is associated with the network interface 30 of the local routing processor 10, the transfer engine 13 sets the packet in queue for the corresponding port 40.

The transfer route searched by the routing

processor 10 is not necessarily single, but the packets can be cast to a plurality of routes at a time. In such a case, the packets can be set in queue for an appropriate buffer of each of the plurality of the
5 routes.

Now, the configuration and the operation of the routing processor will be explained in detail. First, each memory will be explained. Fig. 4 is a diagram for explaining the packet buffer 12 and the
10 header RAM 11.

Fig. 4A shows an example of the format of the packet stored in the packet buffer 12. The packet buffer 12 is supplied with packets from the network 50 or the switch 20. The packet format is that of an IP packet, for example, to which a layer-2 MAC header 401
15 is added. The IP packet includes, for example, a layer-3 IP header 402, a layer-4 header 403 and a payload 404.

The layer-2 MAC header 401 includes a source MAC address (SAMAC) constituting the physical address
20 (hardware address) of the router which has sent the packet immediately before and a destination MAC address (DAMAC) constituting the physical address of the next router to receive the packet. The layer-3 IP header 402 includes a source ID address (hereinafter referred to as the SIP) constituting a source address (address of the
25 transmission terminal) and a destination IP address (hereinafter referred to as the DIP) constituting a destination address (address of the receiving terminal).

The layer-4 header 403 includes a source port
(hereinafter referred to as the SPORT) indicating a
protocol (upper-level application) and a destination
port (hereinafter referred to as the DPORT). The
5 payload 404 includes the user data. In addition, each
header may store the TOS (type of service) indicating
the order of priority and the information such as the
protocol in the upper-level of the IP protocol. These
information can be processed in the same manner as the
10 information described above.

Also, Fig. 4B shows an example format of the
header information stored in the header RAM. The header
information is configured with, for example, the layer-2
MAC header 401 and the layer-3 IP header 402 in the
15 packet format to which the internal header 405 is added
as the control information. The internal header 405
includes the input line number, the output line number
and the QoS control information. The internal packet
format in the router is configured with the packet
20 format of the network to which the internal header 405
is added. In the process, the internal packet can be
formed of the information stored in the packet buffer 12
and the information stored in the header RAM 11. Also,
the internal packet may be transferred from the infor-
25 mation of the packet buffer 12 alone by storing the
internal packet format including the internal header 405
in the packet buffer 12. The transfer control infor-
mation such as the destination information and the

action information searched by the search engine 14 can be written in the internal header 405.

Fig. 5 is a diagram for explaining each table used for route search.

5 As shown in Fig. 5A, the entries in the route table 15 include, for example, the destination IP address 501, the IP address 502 of the next router, the local router transmission RP number 503 and the transmission port number 504. Also, as shown in Fig. 5B, the
10 entries in the ARP table 16 include the IP address 502 of the next router and the MAC address 506 of the next router. Further, as shown in Fig. 5C, the entries of the filter/QoS table 17 include, for example, the value (range) 507 of the IP header/layer-4 header and the
15 action 508. The action 508 includes the filtering process for passing or discarding a packet, the tunneling process for encapsulating or not encapsulating a packet and QoS. Especially, QoS will be explained later again.

20 Fig. 6 is a diagram for explaining the high-speed processing of the routing processor. With reference to this diagram, a method of realizing the packet transfer capable of following a high line speed on the order of gigabits. The high speed is realized by
25 parallel processing or pipelining of the routing. Now, the operation will be explained with reference to the format shown in Figs. 4 and 5.

The routing process is divided roughly into

the receiving process ①, the input search process ②, the output search process ③ and the transmission process ④.

First, in the receiving process ①, the
5 transfer engine 13 receives a packet from the network interface 30. The packet buffer 12 has stored therein an input packet or a packet of the internal packet format with the internal header added thereto. Also, the internal header 405 is added to the layer-2 MAC
10 header 401 and the layer-3 IP header 402 of the input packet to form the header information, which is stored in the RAM 11. The header RAM 11 can be read from and written into at high speed independently of the packet buffer 12, and by storing only the header information
15 therein, the storage capacity can be reduced for further increasing the processing rate. The search engine 14 can access the extracted header information at appropriate timing.

Then, in the input search process ②, the
20 search engine 14 extracts the destination IP address in the layer-3 IP header 402 from the header information, and based on this address, refers to the route table 15 to search for the IP address 502 of the next router, the transmission RP number 503 of the local router and the
25 transmission port number 504. Further, the search engine 14, based on the reference information of the layer-3 IP header 402 and the layer-4 header 403, searches the various items of the action 50 such as QoS

on the input side from the received header information with reference to the filter/QoS table 17. These input-side filter/QoS search and route search are can be executed concurrently since the tables are independently
5 prepared.

Then, in the output search process ③, the search engine 14 extracts the IP address of the next router determined in the input search process ②, and based on this address, searches for the MAC address 506
10 of the next router with reference to the ARP table 16, while at the same time searching for various items of the action 508 on the output side such as QoS with reference to the filter/QoS table 17. The output filter/QoS search and the line table/ARP search can be
15 concurrently executed since each table is prepared independently. The transfer control information including the destination information such as the number/address information of the next destination determined and the action information such as the QoS
20 control information are stored in the header RAM 11. These information can be written, for example, in the internal header 405 or at another appropriate position in the header information.

Then, in the transmission process ④, the
25 header information including the transfer control information searched in the output search process ③ is read from the header RAM 11, and based on the header information and the packet buffer 12, an output packet

is produced and set in queue for the buffer of the network interface 30, another routing processor 10 or the routing manager 60.

Fig. 7 shows an example of the configuration
5 of a search engine in hardware.

The search engine 14 can search the tables including the route table 15, the ARP table 16 and the filter/QoS table 17 for the required data by a tree structure, for example. Now, an explanation will be
10 given of the route search processor for searching for a destination route using the route table 15 as an example of the processor of the search engine 14 configured in hardware.

The route search processor 213 includes a tree
15 structured search circuit 2130, a read address generating circuit 2131 and a route search processing control circuit 2132. The tree structured search circuit 2130 searches the tree structure of 2^p stored in each table such as the route table 15 to generate the
20 pointer of the node next to be read, extract the check bit of the destination IP address of the received packet, determine the end of the tree structure search and update the candidate for the route information resulting from the search. The read address generating
25 circuit 2131 generates the memory address of a part of the words of the node actually read, in accordance with the check bit value and the pointer to the node to be read output from the tree structure search circuit 2130.

The route search processing control circuit 2132, on the other hand, controls the route search processor 213 as a whole (the operation timing and the operating condition of each circuit).

5 Now, the operation of the route search processor 213 will be explained.

 The tree structure search circuit 2130 receives the destination IP address of the received packet from the header RAM 11, and based on this
10 destination IP address and the node mask length, generates the pointer to the next node and delivers it to the read address generating circuit 2131. Also, the tree structured search circuit 2130 extracts the value of the check bit position (check bit value) of the
15 destination IP address indicated by the node mask length and delivers it to the read address generating circuit 2131.

 The read address generating circuit 2131 generates a memory address where the node data to be
20 read is stored, using the pointer to the node, the check bit value and the timing signal from the route search processing control circuit 2132, and transmits it to the memory control circuit 2132. The memory control circuit 2132 generates a memory control signal using the memory
25 address and the timing signal from the route search processing control circuit 2132 and transfers it to the route table 15. The route table 15 that has received this memory control signal transfers a corresponding

node data to the tree structured search circuit 2130 using the signal line 215.

The tree structured search circuit 2130 makes a search using this node data and in the case where it is determined to end the tree structured search, outputs a tree structured search end signal to the route search processing control circuit 2132. The route search processing control circuit 2132 checks a flag with entry in the route information held in the tree structured search circuit 2130, and in the case where the value of the flag is 0, ends the route search process and notifies the transfer engine 13 of the absence of the search result. In the case where the value of the flag with entry is 1, on the other hand, the route information is output to end the search process and the next packet processing is controlled.

Now, Fig. 8 is a diagram for explaining the high-speed processing by the pipelining control. As shown, the receiving process ①, the input search ②, the output search process ③ and the transfer process ④ are carried out by pipelining and thus controlled so that each processor is in constant operation for increasing the speed of the routing process. In the case under consideration, further, the input filter process (input filter/QoS search) and the route table search (route search) are executed in parallel in the input search ②. Also, the output search process ③, the output filter process (output filter/QoS search) and the output line

table search (output line table/ARP search) are executed in parallel. The pipelining is not limited to the structure shown in Fig. 8 but can be implemented in an appropriate sequence.

5 In the pipelining process, upon completion of the first process of the entry N by the processor 1 of all the processors described above, the processor 1 starts the process on the entry N+1 regardless of whether the second process of the entry is completed by
10 the processor 2 for executing the second process subsequent to the first process. This pipelining process can handle N entries in one procession session and therefore the processing speed is quadrupled. In the case described above, the flow search is processed
15 by pipelining divided into four processes. If the process is divided into P processes for pipelining, on the other hand, the performance will be improved by a factor of P.

Fig. 9 is a diagram for explaining the flow
20 search process.

Generally, a network relaying apparatus such as a router lacks a preset connection, and therefore has no connection information unlike in the ATM switch nor QoS control information in the connection information
25 table (packet type communication). As a result, for the router to perform the QoS control, the flow search means for searching the QoS control information with the information in the header is required for each input

packet in addition to the priority transfer function like the ATM switch. As an example, as described below, the priority transfer function is applied to the searched QoS control information by the flow search
5 means. In this case, the conditions for identifying the packets produced by combining such information as the internal information of the header are called the flow conditions, a series of traffic coincident with the flow conditions is called the flow, and to determine whether
10 the input packet meets the flow conditions and to detect the QoS control information and the action information such as transferability information is called the flow search.

According to this embodiment, the QoS control
15 is inserted in the routing processor 10-1 on input side and the routing processor 10-2 on output side, and also the QoS function is provided to the switch 20. The routing processor 10-1 on input side has an input search flow including a filter flow search 911, a tunnel flow
20 search 912 and a QoS flow search 913. In similar fashion, the routing processor 10-2 on output side has an output search flow including a filter flow search 921, a tunnel flow search 922 and a QoS flow search 923. The switch 20 has the arbitration function for selecting
25 the order of transmission according to the priority thus providing the QoS function. The switch 20, like the routing processors 10-1 and 10-2, can be provided with the filter flow search, the tunnel flow search and the

QoS flow search.

The filter flow search 911, 921 determines whether the packet is passed or discarded. The tunnel flow search 912, 922 determines whether the packet is
5 encapsulated or not, and in the case where it is encapsulated, executes the encapsulation software.

The QoS flow search 913, 923 includes the packet priority control, the packet discard control and the band control, for example. The priority control is
10 the one for transmitting the data of high importance degree or data of the real time system in priority. The discard control is the one for discarding the data of low importance degree in the case of heavy traffic or a fault for preventing the loss of important data. The
15 band control, on the other hand, is for segmenting a line into a plurality of bands or changing the bandwidth. For example, the priority control and discard control can be accomplished by controlling the traffic using the matrix of priority class and discard
20 class. In such a case, according to the priority class, the HNA/SNA (Hitachi network architecture/Systems network architecture), voice and animation can be controlled to small delay, while FTP (file transfer protocol), mail and WWW web can be controlled to large
25 delay. According to the discard class, on the other hand, a small discard rate can be set for the control packets and a large discard rate for the voice and animation.

Now, the QoS control by the switch 20 will be explained. The packet sent from the routing processors 10 contains the QoS control information in the control information. The switch 20, especially on output side, carries out the priority control using the QoS control information. Actually, however, this can be accomplished by the output control by setting in queue in the order of priority. As a result, the communication and transfer of an even higher quality is made possible.

Fig. 10 is a diagram for explaining the flow search table.

This flow search table corresponds to the filter/QoS table 17 described above. As an example, as shown in Fig. 10, a reference field 101 includes the source IP address, the destination IP address, the packet length, the IP priority, the IP host protocol, the arrival check flag, the transfer destination TCP/UDP port and the final destination TCP/UDP port. An action field 102, on the other hand, stores therein a filter (pass/discard), a tunnel (encapsulate/not encapsulate) and QoS (delay class, discard class, band, etc.).

Now, a specific method of QoS flow search will be explained. Take the QoS flow search as an example. A similar method can be employed also for the filter flow search or the tunnel flow search. The control information of the respective flows can be stored in mixture in the action field 102, or a flow search table can be prepared for each flow.

First, the linear search method will be explained. In this method, when determining the QoS control information as one action, the preset entries are read sequentially top down from the entry table, and
5 then it is determined whether the values of the header of the packet are all coincident with the valid flow conditions in the reference field 101. In the case of coincidence, the QoS control information in the action field 102 in the entry is determined as the packet QoS
10 control information and the QoS flow search is ended. Once the coincidence with the flow conditions is searched for successfully, the QoS control information in the action field 102 is determined as the QoS control information so that the flow search is ended without
15 executing the next entry search.

In the linear search method described above, it may be difficult to execute the QoS control or filtering at high speed in the network in which a large amount of entries are set. In view of this, the flow
20 search method according to this embodiment desirably employs an input line limiting method or the like in which the flow search can be carried out more rapidly than in the linear search method even in the case where a large amount of entries are set. The input line
25 limiting method will be explained briefly below. In the input line limiting method, only the entries coincident with the input line number making up the reference field of the linear search method are searched to assure high

speed.

Fig. 11 is a diagram for explaining a first input line limiting method. In the first input line limiting method, an entry 511-i with the input line number and the input line number valid bit deleted from the reference field of the linear search method is set for each input line. The flow condition unit 521-i includes the SIP upper limit 501, the SIP lower limit 502, the DIP upper limit 503 and the DIP lower limit 504 indicating the condition for identifying the source or destination user, an IP validity bit 562 indicating the validity of the upper limits and the lower limits of SIP and DIP, the SPORT 505 providing a source port, the DPORT 506 providing a destination port, and a port validity bit 563 indicating the validity of the SPORT 505 and the DPORT 506. The QoS control information unit 530-i includes, for example, the QoS control information 507 used for the priority transfer function. Only the entry 511-i having a coincident input line number providing the flow condition is searched, and therefore the input line number is not required in the entry 511-i. At the time of flow search, only the entry 511-i with the input line thereof assigned a packet is searched.

According to the first input line limiting method described above, assuming that the entry 511-i not related to the input line number is set (set, for example, as "the traffic of Telnet input from all the

input lines is given high priority"), the entries 511-i in the same number as the input line number (= N) are required to be set sometimes leading to a deteriorated efficiency of the memory for realizing the entry table.

5 In view of this, an explanation will be given below of an input line limiting method of higher speed.

Fig. 12 is a diagram for explaining a second input line limiting method. In the second input line limiting method, the lists 540 constituting the
10 addresses in the entry table 750 is set in the list table 760 for each input line. For example, the list 540-11 having the list table address "1" is the address of the entry 511-1, and the list 540-12 having the list table address "2" is the address of the entry 511-H. At
15 the time of flow search, only the list 540 assigned to the input line supplied with a packet is read, and the entry 511-i pointed to by this list 540 is read out. The memory for implementing an entry table can be effectively used if a list 540 having a small bit width
20 (for example, about 10 bits for as many as 1024 entries) is held for each input line and an entry 511-i having a large bit width is shared by the input lines. As a result, a multiplicity of entries 511-i can be set while realizing a high speed operation at the same time.

25 Another example of the flow detection method is the output line limiting method. In the output line limiting method, only the entry 511-i for which the output line number providing the flow condition is

coincident is processed in the same manner as in the input line limiting method described above for realizing a high-speed flow detection. A SAMAC limiting method is available which uses SAMAC instead of the input line
5 number in the header information as the flow condition. In the SAMAC limiting method, the SAMAC group is defined and the entry is limited by the SAMAC identifier providing a SAMAC group identifier, so that the flow search similar to the input line limiting method can be
10 executed.

As will be understood from the foregoing description, according to this invention, there is provided a network relaying apparatus and method for routing packets at high speed while assuring a high
15 communication quality (QoS), a high reliability and security. Also, according to this invention, the hardware processing is carried out for each function block including the transfer engine and the search engine thereby to accomplish a high-speed packet routing
20 and packet transfer. Further, according to the invention, a plurality of tables accessible independently of each other are provided. Also, the routing process is divided into the receiving process, the transmission process, the input search process and the output search
25 process, so that the required tables are used independently for attaining a high speed routing. Further, according to the invention, a still higher processing

speed is realized by executing each process by
pipelining.